

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 09/769,7440

Source: 1Fw16

Date Processed by STIC: 9/23/05

# ***ENTERED***

**CRT Errors Edited by the STIC Systems  
Branch**

Serial Number: 17692442

CRT Edit Date: 9/23/05

Edited by: [signature]

Realigned nucleic acid/amino acid numbers/text in cases where the sequence  
text "wrapped" to the next line

Corrected the SEQ ID NO. Sequence numbers edited were:

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID  
NO's edited:

Deleted: ✓ invalid beginning/end-of-file text ; ✓ page numbers

Inserted mandatory headings/numeric identifiers, specifically:

Moved responses to same line as heading/numeric identifier, specifically:

Other:

deleted alphabetical headings on 1207, 1307, 1407,  
and 1417 lines

Revised 09/09/2003



IFW16

## RAW SEQUENCE LISTING

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:58:35

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

C--> 3 <110> APPLICANT: Microbial Technics Limited  
 4 Le Page, Richard WF  
 5 Wells, Jeremy M  
 6 Hanniffy, Sean B  
 7 Hansbro, Philip M  
 9 <120> TITLE OF INVENTION: Proteins  
 11 <130> FILE REFERENCE: PWC/P21122WO  
 13 <140> CURRENT APPLICATION NUMBER: US/09/769,744D  
 14 <141> CURRENT FILING DATE: 2001-01-26  
 16 <150> PRIOR APPLICATION NUMBER: PCT/GB99/02452  
 17 <151> PRIOR FILING DATE: 1999-07-27  
 19 <150> PRIOR APPLICATION NUMBER: GB 9816336.3  
 20 <151> PRIOR FILING DATE: 1998-07-27  
 22 <150> PRIOR APPLICATION NUMBER: US 60/125329  
 23 <151> PRIOR FILING DATE: 1999-03-19  
 25 <160> NUMBER OF SEQ ID NOS: 196  
 27 <170> SOFTWARE: PatentIn Ver. 2.1  
 29 <210> SEQ ID NO: 1  
 30 <211> LENGTH: 1200  
 31 <212> TYPE: DNA  
 32 <213> ORGANISM: Streptococcus pneumoniae  
 34 <400> SEQUENCE: 1  
 35 atgagaaata tgtgggttgt aatcaaggaa acctatcttc gacatgtcga gtcattggagt 60  
 36 ttcttcttta tgggtgatttc gccgttcctc tttttaggaa tctctgtagg aattgggcat 120  
 37 ctccaagggtt cttctatggc taaaaataat aaagtggcag tagtgacaac agtgccatct 180  
 38 gtagcagaag gactgaagaa tgtaaatggt gtttaacttcg actataaaga cgaagcaagt 240  
 39 gccaaaagaag caattaaaga agaaaaatta aaaggttatt tgaccattga tcaagaagat 300  
 40 agtgttctaa aggcagttta tcatggcgaa acatcgcttg aaaatggaat taaatttgag 360  
 41 gttacaggta cactcaatga actgcaaaat cagcttaatc gttcaactgc ttccttgtct 420  
 42 caagagcagg aaaaacgctt agcgagaca attcaattca cagaaaagat tgatgaagcc 480  
 43 aaggaaaata aaaagtttat tcaaacatt gcagcagggt ccttaggatt ctttctttat 540  
 44 atgattctga ttacctatgc ggggtgtaaca gctcaggaag ttgccagtga aaaaggcacc 600  
 45 aaaattatgg aagtcgtttt ttctagcata agggcaagtc actatttcta tgcgcggatg 660  
 46 atggctctgt ttctagtgat tttaacgcat attgggatct atgttgtagg tgggtctggct 720  
 47 gccgttttgc tctttaaaga tttgccattc ttggctcagt ctggtatttt ggatcacttg 780  
 48 ggagatgcta tctcactgaa taccttgctc tttattttga tcagtctttt catgtacgta 840  
 49 gtcttggcag ccttcctagg atctatgggt tctcgtcctg aggactcagg gaaagccttg 900  
 50 tcgcctttga tgattttgat tatgggtggt ttttttggag tgacagctct aggtgcagct 960  
 51 ggtgacaatc tcctcttgaa gattgggtct tatattccct ttatttcgac cttctttatg 1020  
 52 ccgtttcgaa cgattaatga ctatgcgggg ggagcagaag catggatttc acttgctatt 1080  
 53 acagtgattt ttgcgggtgg agcaacagga tttatcggac gcatgtatgc tagtctcggt 1140  
 54 cttcaaacgg atgatttagg gatttggaac acctttaaac gtgccttatc ttataaatag 1200  
 57 <210> SEQ ID NO: 2

P.6

## RAW SEQUENCE LISTING

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:58:35

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

```

58 <211> LENGTH: 399
59 <212> TYPE: PRT
60 <213> ORGANISM: Streptococcus pneumoniae
62 <400> SEQUENCE: 2
63 Met Arg Asn Met Trp Val Val Ile Lys Glu Thr Tyr Leu Arg His Val
64   1           5           10           15
66 Glu Ser Trp Ser Phe Phe Phe Met Val Ile Ser Pro Phe Leu Phe Leu
67           20           25           30
69 Gly Ile Ser Val Gly Ile Gly His Leu Gln Gly Ser Ser Met Ala Lys
70           35           40           45
72 Asn Asn Lys Val Ala Val Val Thr Thr Val Pro Ser Val Ala Glu Gly
73           50           55           60
75 Leu Lys Asn Val Asn Gly Val Asn Phe Asp Tyr Lys Asp Glu Ala Ser
76  65           70           75           80
78 Ala Lys Glu Ala Ile Lys Glu Glu Lys Leu Lys Gly Tyr Leu Thr Ile
79           85           90           95
81 Asp Gln Glu Asp Ser Val Leu Lys Ala Val Tyr His Gly Glu Thr Ser
82           100          105          110
84 Leu Glu Asn Gly Ile Lys Phe Glu Val Thr Gly Thr Leu Asn Glu Leu
85           115          120          125
87 Gln Asn Gln Leu Asn Arg Ser Thr Ala Ser Leu Ser Gln Glu Gln Glu
88           130          135          140
90 Lys Arg Leu Ala Gln Thr Ile Gln Phe Thr Glu Lys Ile Asp Glu Ala
91 145          150          155          160
93 Lys Glu Asn Lys Lys Phe Ile Gln Thr Ile Ala Ala Gly Ala Leu Gly
94           165          170          175
96 Phe Phe Leu Tyr Met Ile Leu Ile Thr Tyr Ala Gly Val Thr Ala Gln
97           180          185          190
99 Glu Val Ala Ser Glu Lys Gly Thr Lys Ile Met Glu Val Val Phe Ser
100          195          200          205
102 Ser Ile Arg Ala Ser His Tyr Phe Tyr Ala Arg Met Met Ala Leu Phe
103          210          215          220
105 Leu Val Ile Leu Thr His Ile Gly Ile Tyr Val Val Gly Gly Leu Ala
106 225          230          235          240
108 Ala Val Leu Leu Phe Lys Asp Leu Pro Phe Leu Ala Gln Ser Gly Ile
109           245          250          255
111 Leu Asp His Leu Gly Asp Ala Ile Ser Leu Asn Thr Leu Leu Phe Ile
112           260          265          270
114 Leu Ile Ser Leu Phe Met Tyr Val Val Leu Ala Ala Phe Leu Gly Ser
115           275          280          285
117 Met Val Ser Arg Pro Glu Asp Ser Gly Lys Ala Leu Ser Pro Leu Met
118          290          295          300
120 Ile Leu Ile Met Gly Gly Phe Phe Gly Val Thr Ala Leu Gly Ala Ala
121 305          310          315          320
123 Gly Asp Asn Leu Leu Lys Ile Gly Ser Tyr Ile Pro Phe Ile Ser
124           325          330          335
126 Thr Phe Phe Met Pro Phe Arg Thr Ile Asn Asp Tyr Ala Gly Gly Ala
127           340          345          350
129 Glu Ala Trp Ile Ser Leu Ala Ile Thr Val Ile Phe Ala Val Val Ala

```

## RAW SEQUENCE LISTING

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:58:35

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

```

130          355          360          365
132 Thr Gly Phe Ile Gly Arg Met Tyr Ala Ser Leu Val Leu Gln Thr Asp
133          370          375          380
135 Asp Leu Gly Ile Trp Lys Thr Phe Lys Arg Ala Leu Ser Tyr Lys
136 385          390          395
140 <210> SEQ ID NO: 3
141 <211> LENGTH: 1125
142 <212> TYPE: DNA
143 <213> ORGANISM: Streptococcus pneumoniae
145 <400> SEQUENCE: 3
146 cctgggaaaag tcttgaaaat tatgatagaa tgggtggaagg aaaaattcag gagagtagta 60
147 gtgactcaaaa atgttgaaaag tcttctcgta tccattgttaa tcagtgcata caatgaagaa 120
148 aaatatctgc ctggtctaata tgaagactta aaaaatcaaaa cctatcctaa agaggatatt 180
149 gaaattctat ttataaatgc tatgtccaca gatgggacca cagctatcat tcagcaattt 240
150 ataaaggaag atacagagtt taactcaatt agattgtata acaatcctaa gaaaaatcaa 300
151 gctagtgggt ttaacctggg agttaaacat tctgtagggg accttatttt aaaaattgat 360
152 gctcattcaa aagttactga gacttttgta atgaacaatg tggctattat tcaacaagg 420
153 gaatttgtct gtgggggggcc tagaccgacg attgtcgaag gaaaaggaaa atgggagag 480
154 accttgcatc ttgttgagga aaatatgttt ggcagtagca ttgccaatta tcgaaatagt 540
155 tctgaggata gatatgtttc ttctattttt catggaatgt ataaacgaga ggttttccag 600
156 aagggttggt tagtaaatga gcaacttggc cgaactgaag ataatgatat tcattataga 660
157 attcgagaat atggttataa aatccgctat agcccaagta ttctatctta tcagtatatt 720
158 cgaccaacat tcaagaaaat gctgcacaa aagtattcaa atggtttgtg gattggcttg 780
159 acaagtcacg ttcagtttaa gtgtttatca ttatttcaat atgttccttg tttatttggt 840
160 ttgagtcctg tgttttagtct agcattgtta ccgatcacat tcgtattcat aactttacta 900
161 ttaggtgcct attttctact tttgtcatta ctactttgc tgactttatt aaaacataaa 960
162 aatggatttc taattgtgat gccctttatt ttattttcca ttcactttgc ttatggcctt 1020
163 gggacgattg taggtttaat tagaggattt aaatggaaga aggagtacaa gagaacaata 1080
164 atttatttgg ataaaaataa ccaaataaat caaaatatgc tataa 1125
167 <210> SEQ ID NO: 4
168 <211> LENGTH: 374
169 <212> TYPE: PRT
170 <213> ORGANISM: Streptococcus pneumoniae
172 <400> SEQUENCE: 4
173 Pro Gly Lys Val Leu Lys Ile Met Ile Glu Trp Trp Lys Glu Lys Phe
174 1 5 10 15
176 Arg Arg Val Val Val Thr Gln Asn Val Glu Ser Leu Leu Val Ser Ile
177 20 25 30
179 Val Ile Ser Ala Tyr Asn Glu Glu Lys Tyr Leu Pro Gly Leu Ile Glu
180 35 40 45
182 Asp Leu Lys Asn Gln Thr Tyr Pro Lys Glu Asp Ile Glu Ile Leu Phe
183 50 55 60
185 Ile Asn Ala Met Ser Thr Asp Gly Thr Thr Ala Ile Ile Gln Gln Phe
186 65 70 75 80
188 Ile Lys Glu Asp Thr Glu Phe Asn Ser Ile Arg Leu Tyr Asn Asn Pro
189 85 90 95
191 Lys Lys Asn Gln Ala Ser Gly Phe Asn Leu Gly Val Lys His Ser Val
192 100 105 110
194 Gly Asp Leu Ile Leu Lys Ile Asp Ala His Ser Lys Val Thr Glu Thr

```

## RAW SEQUENCE LISTING

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:58:35

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

```

195          115          120          125
197 Phe Val Met Asn Asn Val Ala Ile Ile Gln Gln Gly Glu Phe Val Cys
198          130          135          140
200 Gly Gly Pro Arg Pro Thr Ile Val Glu Gly Lys Gly Lys Trp Ala Glu
201 145          150          155          160
203 Thr Leu His Leu Val Glu Glu Asn Met Phe Gly Ser Ser Ile Ala Asn
204          165          170          175
206 Tyr Arg Asn Ser Ser Glu Asp Arg Tyr Val Ser Ser Ile Phe His Gly
207          180          185          190
209 Met Tyr Lys Arg Glu Val Phe Gln Lys Val Gly Leu Val Asn Glu Gln
210          195          200          205
212 Leu Gly Arg Thr Glu Asp Asn Asp Ile His Tyr Arg Ile Arg Glu Tyr
213          210          215          220
215 Gly Tyr Lys Ile Arg Tyr Ser Pro Ser Ile Leu Ser Tyr Gln Tyr Ile
216 225          230          235          240
218 Arg Pro Thr Phe Lys Lys Met Leu His Gln Lys Tyr Ser Asn Gly Leu
219          245          250          255
221 Trp Ile Gly Leu Thr Ser His Val Gln Phe Lys Cys Leu Ser Leu Phe
222          260          265          270
224 His Tyr Val Pro Cys Leu Phe Val Leu Ser Leu Val Phe Ser Leu Ala
225          275          280          285
227 Leu Leu Pro Ile Thr Phe Val Phe Ile Thr Leu Leu Leu Gly Ala Tyr
228          290          295          300
230 Phe Leu Leu Leu Ser Leu Leu Thr Leu Leu Thr Leu Leu Lys His Lys
231 305          310          315          320
233 Asn Gly Phe Leu Ile Val Met Pro Phe Ile Leu Phe Ser Ile His Phe
234          325          330          335
236 Ala Tyr Gly Leu Gly Thr Ile Val Gly Leu Ile Arg Gly Phe Lys Trp
237          340          345          350
239 Lys Lys Glu Tyr Lys Arg Thr Ile Ile Tyr Leu Asp Lys Ile Ser Gln
240          355          360          365
242 Ile Asn Gln Asn Met Leu
243          370
247 <210> SEQ ID NO: 5
248 <211> LENGTH: 696
249 <212> TYPE: DNA
250 <213> ORGANISM: Streptococcus pneumoniae
252 <400> SEQUENCE: 5
253 atgatgaaag acaaaaatac gatagaaatc gatgtatttc aattagttaa aagcttgtgg 60
254 aaacgcaagc taatgatttt aatagtggca cttgtgacag gtgcgggggc ttttgcata 120
255 agcactttta ttgttaagcc agaataacg agtaccacgc gaatttacgt agtgaatcgc 180
256 aatcaaggag acaagccggg gttgacaaat caggatttgc aggcaggaac ttatctggta 240
257 aaagactacc gtgagattat cctttcgcag gatgttttgg aggaagttgt ttctgatttg 300
258 aaactagatt tgacgcaaaa aggtttggct aataaaatta aagtgacagt accagttgat 360
259 acccgatttg tctctatttc agttaatgat cgagttcctg aagaggcaag ccgtatcgct 420
260 aactctttga gagaagtagc tgctcaaaaa attatcagta ttactcgtgt ttctgacgtg 480
261 acaacactgg aggaggcaag gccggcgata tccccgtctt cgccaaatat taaacgcaat 540
262 acactaattg gttttttggc aggggtgatt ggaactagtg ttatagttct tcatcttgaa 600
263 cttttggata ctctgttgaa acgtccggaa gatatcgaaa atacattgca gatgacactt 660

```

## RAW SEQUENCE LISTING

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:58:35

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

```

264 ttgggagttg tgccaaactt gggtaagttg aaatag                                696
267 <210> SEQ ID NO: 6
268 <211> LENGTH: 231
269 <212> TYPE: PRT
270 <213> ORGANISM: Streptococcus pneumoniae
272 <400> SEQUENCE: 6
273 Met Met Lys Glu Gln Asn Thr Ile Glu Ile Asp Val Phe Gln Leu Val
274   1           5           10           15
276 Lys Ser Leu Trp Lys Arg Lys Leu Met Ile Leu Ile Val Ala Leu Val
277           20           25           30
279 Thr Gly Ala Gly Ala Phe Ala Tyr Ser Thr Phe Ile Val Lys Pro Glu
280           35           40           45
282 Tyr Thr Ser Thr Thr Arg Ile Tyr Val Val Asn Arg Asn Gln Gly Asp
283           50           55           60
285 Lys Pro Gly Leu Thr Asn Gln Asp Leu Gln Ala Gly Thr Tyr Leu Val
286           65           70           75           80
288 Lys Asp Tyr Arg Glu Ile Ile Leu Ser Gln Asp Val Leu Glu Glu Val
289           85           90           95
291 Val Ser Asp Leu Lys Leu Asp Leu Thr Pro Lys Gly Leu Ala Asn Lys
292           100          105          110
294 Ile Lys Val Thr Val Pro Val Asp Thr Arg Ile Val Ser Ile Ser Val
295           115          120          125
297 Asn Asp Arg Val Pro Glu Glu Ala Ser Arg Ile Ala Asn Ser Leu Arg
298           130          135          140
300 Glu Val Ala Ala Gln Lys Ile Ile Ser Ile Thr Arg Val Ser Asp Val
301          145          150          155          160
303 Thr Thr Leu Glu Glu Ala Arg Pro Ala Ile Ser Pro Ser Ser Pro Asn
304           165          170          175
306 Ile Lys Arg Asn Thr Leu Ile Gly Phe Leu Ala Gly Val Ile Gly Thr
307           180          185          190
309 Ser Val Ile Val Leu His Leu Glu Leu Leu Asp Thr Arg Val Lys Arg
310           195          200          205
312 Pro Glu Asp Ile Glu Asn Thr Leu Gln Met Thr Leu Leu Gly Val Val
313           210          215          220
315 Pro Asn Leu Gly Lys Leu Lys
316          225          230
320 <210> SEQ ID NO: 7
321 <211> LENGTH: 555
322 <212> TYPE: DNA
323 <213> ORGANISM: Streptococcus pneumoniae
325 <400> SEQUENCE: 7
326 atggtaaaag tagcagttat attagctcag ggctttgaag aaattgaagc cttgacagtt 60
327 gtagatgtct tgcgtcgagc caatatcaca tgtgatatgg ttggttttga agagcaagta 120
328 acgggttcgc atgcaatcca agtaagagca gatcatgtct ttgatggaga tttatcagac 180
329 tatgatatga ttgttcttcc tggaggtatg cctggttctg cacatttacg tgataatcag 240
330 accttgattc aagaattgca aagcttcgag caagaaggga agaaactagc agccatttgt 300
331 gcggcaccaa ttgccctcaa tcaagcagag atattgaaaa ataagcgata cacttgttat 360
332 gacggcggtt aagagcaaat ccttgatggt cactacgtca aggaaacagt agtggttagat 420
333 ggtcagttga caaccagtcg gggtccttca acagcccttg cctttgccta cgagttggtg 480

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 09/23/2005  
PATENT APPLICATION:    US/09/769,744D      TIME: 14:58:36

Input Set : N:\AMC\769744.txt  
Output Set: N:\CRF4\09232005\I769744D.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:182; Xaa Pos. 3



**VERIFICATION SUMMARY**

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:58:36

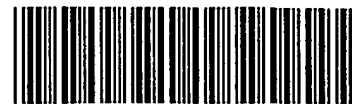
Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:9148 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:182 after pos.:0

**Raw Sequence Listing before editing  
(for reference only)**



IFW16

## RAW SEQUENCE LISTING

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:55:05

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

3 <110> APPLICANT: Microbial Technics Limited  
 4 Le Page, Richard WF  
 5 Wells, Jeremy M  
 6 Hanniffy, Sean B  
 7 Hansbro, Philip M  
 9 <120> TITLE OF INVENTION: Proteins  
 11 <130> FILE REFERENCE: PWC/P21122WO  
 13 <140> CURRENT APPLICATION NUMBER: US/09/769,744D  
 C--> 14 <141> CURRENT FILING DATE: 2001-01-26  
 16 <150> PRIOR APPLICATION NUMBER: PCT/GB99/02452  
 17 <151> PRIOR FILING DATE: 1999-07-27  
 19 <150> PRIOR APPLICATION NUMBER: GB 9816336.3  
 20 <151> PRIOR FILING DATE: 1998-07-27  
 22 <150> PRIOR APPLICATION NUMBER: US 60/125329  
 23 <151> PRIOR FILING DATE: 1999-03-19  
 25 <160> NUMBER OF SEQ ID NOS: 196  
 27 <170> SOFTWARE: PatentIn Ver. 2.1

Does Not Comply  
Corrected Diskette Needed

## ERRORED SEQUENCES

9317 <210> SEQ ID NO: 196  
 9318 <211> LENGTH: 11  
 9319 <212> TYPE: PRT  
 9320 <213> ORGANISM: Streptococcus pneumoniae  
 9322 <400> SEQUENCE: 196  
 9323 Gln Lys Ile Thr Met Ile Thr Phe Thr Phe Gln  
 9324 1 5 10  
 E--> 9330 163

see p. 2

09/769,744D

2

SEQUENCE LISTING

<110> Microbial Technics Limited  
Le Page, Richard WF  
Wells, Jeremy M  
Hanniffy, Sean B  
Hansbro, Philip M

*delete alphabetical headers*

<120> ~~TITLE OF INVENTION~~: Proteins

<130> ~~FILE REFERENCE~~: PWC/P21122WO

<140> ~~CURRENT APPLICATION NUMBER~~: US/09/769,744

<141> ~~CURRENT FILING DATE~~: 2001-02-2662785

**VERIFICATION SUMMARY**

DATE: 09/23/2005

PATENT APPLICATION: US/09/769,744D

TIME: 14:55:07

Input Set : N:\AMC\769744.txt

Output Set: N:\CRF4\09232005\I769744D.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:9148 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:182 after pos.:0

L:9330 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:196